

Digital Object Identifier

# Federated Learning model using CBDPL for Medical Image Segmentation

K. HEMALATHA<sup>1</sup>, SHRADHANJALI DAS<sup>2</sup>

<sup>1</sup>School of Computer Science and Engineering, Vellore Institute of Technology - Chennai Campus, Chennai, Tamil Nadu 600127, India (e-mail: hemalatha.kasirajan@vit.ac.in)

<sup>2</sup>School of Computer Science and Engineering, Vellore Institute of Technology - Chennai Campus, Chennai, Tamil Nadu 600127, India (e-mail: shradhanjali.das2021@vitstudent.ac.in)

Corresponding author: K. Hemalatha (e-mail: hemalatha.kasirajan@vit.ac.in).

**ABSTRACT** Federated learning (FL) allows hospitals and medical centers to train models without sharing patient data. Instead of sending data to a central server, each institution trains the model locally and only shares updates, keeping patient information private. Our study proposes a novel methodology for medical image segmentation that leverages semi-supervised learning and confidence-based dynamic pseudo-labeling (CBDPL) to enhance model performance with limited labeled data. Each participating client site independently performs feature extraction and basic segmentation tasks. For labeled data, segmentation is guided by a ground-truth comparison and loss minimization. For unlabeled data, semi-supervised learning is applied, with a dynamic pseudo-labeling module that assigns pseudo-labels based on confidence scores, integrating only high-confidence predictions to refine the model iteratively. Client updates are then combined on a central server. It helps to improve the overall model while keeping each client's data private. This approach demonstrates the potential for robust, privacy-preserving segmentation in multi-dataset medical image analysis settings, addressing labeled data scarcity and privacy concerns.

**INDEX TERMS** confidence-based dynamic pseudo-labeling, federated learning, medical image segmentation, privacy-preserving framework ,semi-supervised learning

## I. INTRODUCTION

Medical image analysis is important in modern healthcare, it helps early diagnosis, precise treatment planning, and improved prediction for various diseases. Even though deep learning has improved medical image segmentation, it still relies on having a large number of labeled images to work accurately. However, labeling medical images is a resource-intensive process that requires a lot of time and domain expertise of medical professionals. These limitations hinder the development of generalized and robust diagnostic models, particularly for smaller healthcare institutions with limited resources [1]. FL enables different hospitals to train a shared model together without requiring raw data to be shared, thereby ensuring compliance with privacy regulations [23], [6], [29]. In a typical FL setup, each hospital trains a model using its data and only sends updates to a central server, preserving patient confidentiality while facilitating global model development. However, traditional FL frameworks are often limited to labeled datasets, but there is a large availability of unlabeled datasets [22], [11]. Due to the growth of data-efficient learning, many labeled and unlabeled datasets are available. This creates a need for methods that can use both

datasets to improve segmentation while preserving privacy. [27]

Traditional FL methods face multiple limitations, particularly in medical imaging applications. They struggle with: (1) reliance on large labeled datasets, which are often scarce in medical domains, (2) label noise or class imbalance across clients, (3) underutilization of the large volume of unlabeled data, (4) reduced performance due to data heterogeneity across clients, and (5) privacy risks if semi-supervised techniques are not integrated carefully.

To address these limitations, our work proposes several key innovations. First, to reduce dependence on large labeled datasets, we integrate semi-supervised learning (SSL) into the federated framework, enabling effective use of both labeled and unlabeled data. To mitigate label noise and class imbalance across clients, our confidence-based dynamic pseudo-labeling (CBDPL) strategy selectively assigns pseudo-labels based on prediction certainty, reducing the influence of unreliable samples. We also tackle the underutilization of unlabeled data by iteratively refining pseudo-labels, allowing even data without annotations to contribute meaningfully to the learning process. To manage client data

heterogeneity, our framework promotes model generalization through collaborative training, ensuring robust performance across diverse clinical settings. Finally, privacy concerns are addressed by ensuring that no raw data leaves local institutions, and by designing the SSL mechanisms to operate entirely within the privacy-preserving boundaries of FL.

The motivation behind this work stems from the practical constraints in real-world medical scenarios, where labeled data is limited, unlabeled data is abundant, and privacy is a top concern. In this work, we focus on enhancing the federated learning paradigm by proposing a novel method within FL that leverages SSL and a CBDPL strategy for medical image segmentation. It is important to note that our focus is not merely on proposing a specific model, but rather on advancing how medical image segmentation can be conducted collaboratively and efficiently under the constraints and privacy requirements of FL. The CBDPL strategy allows institutions with limited labeled data to contribute and benefit from a shared global model, which improves performance while maintaining patient confidentiality.

We propose a privacy-preserving FL framework that integrates SSL and a novel CBDPL mechanism to address the challenges of label scarcity. Unlike conventional FL methods that rely solely on labeled data, our approach effectively utilizes labeled and unlabeled medical images, allowing for improved segmentation even in resource-constrained settings. The CBDPL mechanism dynamically assigns pseudo-labels to unlabeled data based on prediction confidence, ensuring that only high-confidence pseudo-labels contribute to model refinement. This iterative process progressively enhances segmentation accuracy while mitigating noise from low-confidence predictions. By combining SSL with FL, our framework enables efficient learning across multiple medical institutions while keeping data private.

The following concisely outlines this work's main contributions:

I. We propose a secure FL framework that enables multiple medical institutions to collaboratively train segmentation models without sharing raw patient data, thereby preserving data privacy.

II. We integrate SSL into the FL framework to leverage both labeled and unlabeled medical images, improving segmentation performance, particularly for clients with limited labeled data.

III. We introduce a CBDPL mechanism that selectively assigns high-confidence pseudo-labels to unlabeled data, effectively reducing label noise and enhancing segmentation accuracy in low-resource settings.

The rest of this paper is organized as follows. Section 2 provides a detailed review of related work and contemporary methods in FL, SSL, and medical image segmentation. Section 3 presents the proposed FL framework, including the CBDPL methodology, dynamic pseudo-labeling mechanism, and global model aggregation strategy. Section 4 discusses the experimental results, including performance evaluation

and comparisons with existing techniques. Finally, Section 5 concludes the paper and outlines potential directions for future research.

## II. RELATED WORK

In medical image analysis, FL and SSL have drawn a lot of interest because of the potential to improve model training while preserving data privacy and resolving the lack of labeled datasets. Various studies have explored their applications across different medical imaging tasks, from disease diagnosis to segmentation. This section provides an overview of FL in medical imaging, the integration of SSL in federated environments, advancements in pseudo-labeling techniques, and privacy-preserving methods in FL-based frameworks.

### A. FEDERATED LEARNING IN MEDICAL IMAGING

FL has demonstrated potential as a platform for collaboratively training deep learning models across various healthcare facilities while protecting patient privacy. Yan et al. introduced Variation-Aware Federated Learning (VAFL) to address inter-client variations in medical image analysis tasks, such as prostate cancer classification, by adopting personalized model updates for improved robustness across decentralized data sources [37]. Similarly, Ullah et al. developed a scalable FL approach for medical imaging, ensuring the effective participation of intermittent clients in collaborative smart healthcare systems [28]. Yue et al. proposed a specificity-aware FL framework with dynamic feature fusion networks to handle class imbalances in medical image classification [40].

Other studies have extended FL to handle domain-specific challenges. Similarly, Lei et al. proposed a federated Transformer-based domain adaptation model for Alzheimer's disease diagnosis, enabling effective feature extraction and transfer across multi-site MRI data [13]. Rui et al. tackled data heterogeneity by introducing a label-efficient self-supervised FL framework, showcasing its ability to handle diverse medical imaging datasets [36]. Xu et al. explored federated multi-organ segmentation with inconsistent labels, highlighting the adaptability of FL models in scenarios with label heterogeneity [35]. Furthermore, numerous studies have applied FL to various cancer types, including prostate cancer [21], [3], [12], [17], lung cancer [9], [25], [8], breast cancer [14], [10], and brain cancer [1], demonstrating its utility in privacy-preserving collaborative learning for diagnosis and treatment.

A comprehensive review by Sohan et al. highlighted the advancements in FL for medical image analysis, summarizing key trends, challenges, and promising directions for future research [24]. Zhou et al. also extensively reviewed deep learning in medical imaging, discussing imaging traits, technological trends, and FL's role in advancing healthcare [41].

## B. SEMI-SUPERVISED LEARNING IN FEDERATED ENVIRONMENTS

The scarcity of labeled medical data has driven research towards integrating SSL within federated frameworks [2], [31]. Wu et al. introduced a prototype-based pseudo-labeling method combined with contrastive learning to improve federated semi-supervised segmentation, showcasing superior performance on COVID-19 and other medical datasets [34]. This work highlights the importance of pseudo-labeling strategies for leveraging unlabeled data.

In a similar vein, Qiu et al. proposed a pseudo-label denoising mechanism to mitigate noise in pseudo-labeled data for medical image segmentation, demonstrating significant improvements in federated SSL settings [20]. Chen et al. further explored attention mechanisms in vertical federated learning, leveraging both metadata and image features for more accurate semi-supervised tasks [4]. Jin et al. proposed FedDP, a dual-personalization framework designed for medical image segmentation tasks, effectively addressing challenges in personalization and heterogeneity [30].

## C. PSEUDO-LABELING

Pseudo-labeling techniques based on prediction confidence have shown promise for integrating unlabeled data effectively into the training process. Guo et al. introduced a dual class-aware contrastive framework, wherein confidence-based pseudo-labels guide model updates through local and global class prototypes, achieving superior segmentation performance in semi-supervised federated learning settings [7]. Additionally, Wicaksana et al. proposed FedMix, a mixed supervision strategy that combines adaptive pseudo-labeling with federated learning to enhance segmentation accuracy while balancing labeled and unlabeled data contributions [33]. Sun et al. presented FKD-Med, which leverages knowledge distillation for privacy-aware and communication-efficient FL in medical image segmentation [26].

## D. PRIVACY-PRESERVING TECHNIQUES IN FEDERATED LEARNING

Ensuring data privacy remains a critical aspect of federated learning in medical imaging. Mantey et al. explored homomorphic encryption to secure federated training processes in medical recommender systems, highlighting the feasibility of encrypted computation for sensitive healthcare data [16]. Similarly, Liu et al. presented a comprehensive privacy-preserving framework with secure authentication mechanisms, ensuring robust protection for federated learning systems in Internet of Medical Things (IoMT) environments [15].

In addition to encryption-based approaches, Wang et al. proposed a privacy-preserving split learning model integrated with differential privacy, tailored for large-scale vision pre-training tasks, further extending its utility to medical imaging datasets [32]. Mu et al. introduced explainable federated medical image analysis using causal learning and blockchain, enhancing both privacy and interpretability [18].

## E. EMERGING TRENDS IN PERSONALIZED FEDERATED LEARNING AND LLMS

Recent advances in federated learning have focused on personalization and the integration of large-scale foundation models to improve performance and generalization. A notable example is the work by Xia et al., which introduces hypernetworks to personalize model weights for individual clients based on scanning physics, thereby improving performance across heterogeneous CT datasets. This method provides a client-specific adaptation strategy that aligns with real-world clinical variability and complements conventional FL methods relying on global model aggregation. [39]

Furthermore, large language models (LLMs) are beginning to shape the future of FL in medical imaging. The study by Yang et al. explores how LLMs can extract patient-specific anatomical and scanning-related context to guide federated CT image denoising. This integration of LLMs with FL frameworks introduces a powerful paradigm shift, where models are not only privacy-preserving and semi-supervised but also context-aware and generative. [38]

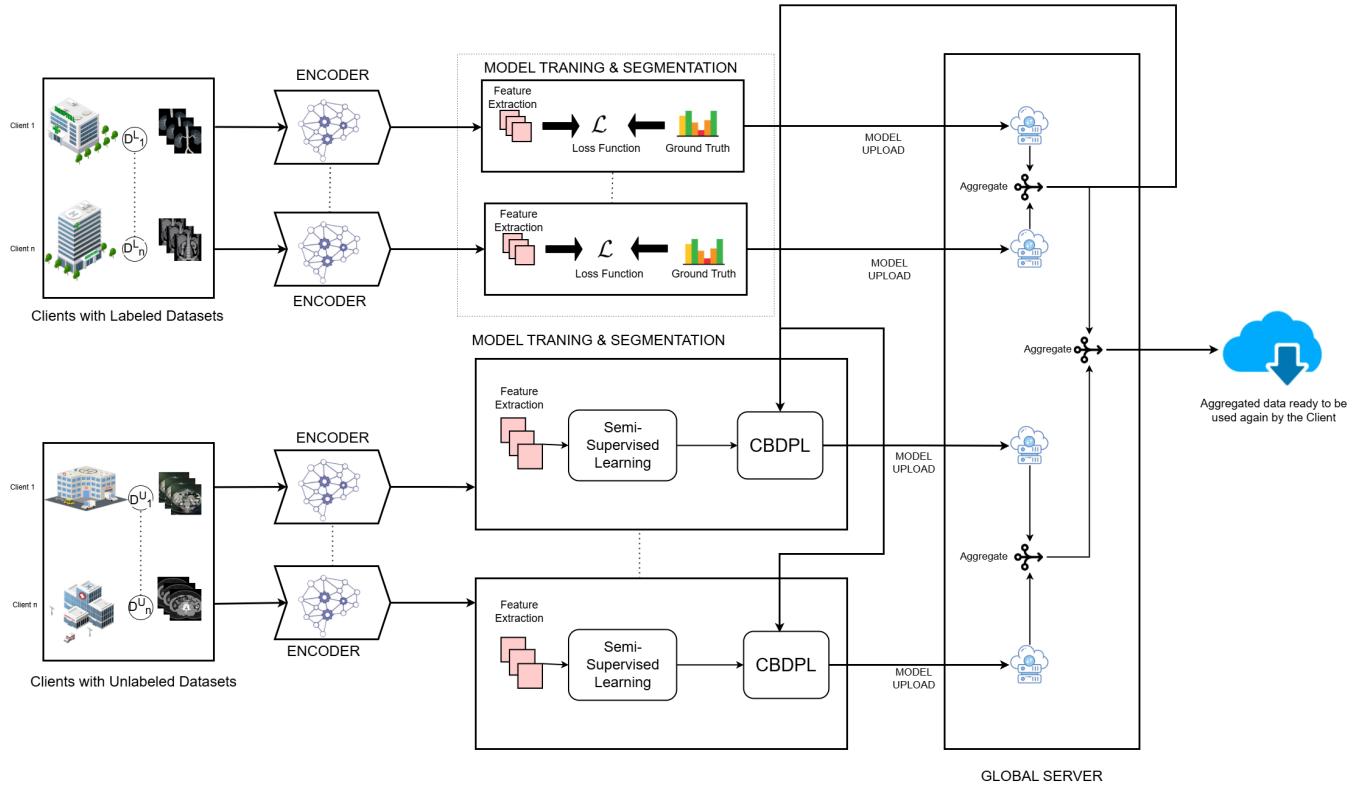
In addition, the work by Yingyu et al. incorporates uncertainty-aware learning into semi-supervised medical image segmentation. This approach uses uncertainty quantification to guide the selection of reliable pseudo-labels for unlabeled data, enhancing the segmentation performance by focusing on regions with higher predictive certainty. The method is highly relevant to federated settings, as it can effectively reduce label noise and improve model performance even with limited labeled data across distributed clients. [5]

These trends emphasize the evolving landscape of FL research, where personalized model learning and multi-modal understanding through LLMs represent the next frontier. Our work is distinct in focusing on confidence-based pseudo-labeling for semi-supervised segmentation across clients with limited labeled data, but future extensions may explore the synergy between pseudo-labeling and LLM-guided personalization.

## F. RESEARCH GAPS AND CONTRIBUTIONS

In previous studies, FL has been explored for medical image segmentation, but challenges remain in handling limited labeled data, noisy pseudo-labels, and maintaining robust privacy. Existing SSL techniques often lack dynamic mechanisms to filter low-confidence pseudo-labels, which may degrade model performance. Our proposed framework addresses these limitations through:

- CBDPL, which iteratively refines pseudo-labels by integrating only high-confidence predictions.
- A robust SSL with FL architecture that leverages both labeled and unlabeled data for segmentation tasks.
- Preservation of data privacy through decentralized training, ensuring no raw medical data is shared across institutions.



**FIGURE 1.** Architecture of FL model using CBDPL for Medical Image Segmentation. Here  $D_n^L$  implies the clients with labeled datasets, while  $D_n^U$  implies the clients with unlabeled datasets.

### III. METHODOLOGY

This work proposes a novel FL framework for medical image segmentation, addressing the challenges of privacy-preserving multi-client collaboration and limited labeled data. The framework begins with data pre-processing and local training on client datasets, where labeled data is utilized for supervised segmentation tasks. In contrast, unlabeled data are processed using an SSL approach. A CBDPL mechanism is applied to assign high-confidence pseudo-labels to unlabeled data, ensuring noise minimization and iterative refinement of the model. Finally, client updates are combined on a central server using federated averaging (FedAvg), enabling the development of a robust global segmentation model that maintains data privacy across all participating institutions.

#### A. DATASET DESCRIPTION

The Multi Cancer Dataset from Kaggle [19] comprises medical images across eight major cancer types and their 26 subclasses. These include Acute Lymphoblastic Leukemia, Breast Cancer, Cervical Cancer, Brain Cancer, Kidney Cancer, Lung and Colon Cancer, Lymphoma, and Oral Cancer. The dataset's diversity and wide-ranging pathology ensure it is well-suited for developing robust deep learning models capable of detecting and classifying multiple types of cancerous cells. Each image in the dataset is a visual representation of pathological samples, annotated and labeled appropriately for supervised learning tasks.

#### B. DATA PREPARATION

FL is particularly well-suited for medical imaging applications, where data privacy and decentralization are critical concerns. This dataset closely mirrors real-world healthcare scenarios where medical institutions maintain localized datasets with privacy constraints. These diverse cancer classes present challenges in feature extraction, model generalization, and class imbalance, making them ideal candidates for SSL approaches.

The dataset was split into labeled (30%) and unlabeled (70%) subsets, distributed across multiple clients  $C = \{C_1, C_2, \dots, C_n\}$ , simulating decentralized healthcare environments. Data pre-processing included image resizing, normalization, and augmentation to improve generalization. Let  $D_i^L = \{(x_j, y_j)\}_{j=1}^m$  represent the labeled dataset of client  $C_i$ , where  $y_j$  is the label for the  $j$ -th image  $x_j$ . Similarly, let  $D_i^U = \{(u_k)\}_{k=1}^n$  denote the unlabeled dataset, where  $u_k$  is the input image with no label.

The preprocessing steps were uniformly applied across all client datasets to enhance model performance and ensure data consistency. All images were resized to a fixed resolution of 224x224x3 to match the input size required by the DenseNet121 model. Pixel intensities were normalized to a range of [0,1] to maintain stability during training. Additionally, data augmentation techniques, including random flips, rotations, cropping, and intensity adjustments, were implemented to improve model generalization and mitigate

Cancer Type	Description	Image Count
Acute Lymphoblastic Leukemia (ALL)	Non-cancerous cells, early leukemia stages, pre-stage abnormal cells, advanced leukemia cells	20,000
Brain Cancer	Common brain tumor, tumors affecting brain membranes, pituitary gland tumors	15,000
Breast Cancer	Non-cancerous breast tissues, cancerous breast tissues	10,000
Cervical Cancer	Abnormal cell growth, viral infection-related changes, precancerous/metaplastic cells, immature/mature squamous cells	25,000
Kidney Cancer	Healthy kidney tissues, tumor-affected kidney tissues	10,000
Lung and Colon Cancer	Colon cancer cells, healthy colon tissues, lung cancer cells, healthy lung tissues, aggressive lung cancer	25,000
Lymphoma	Slow-growing blood cancer, non-Hodgkin lymphoma, aggressive lymphoma	15,000
Oral Cancer	Healthy oral tissues, cancerous oral cells	10,000

TABLE 1. Overview of the Multi-Cancer Dataset.

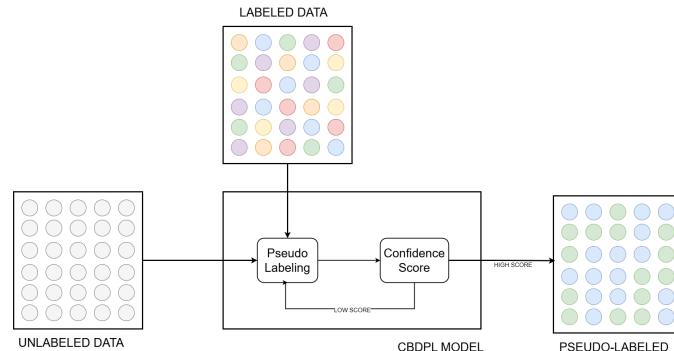


FIGURE 2. Architecture of CBDPL model.

overfitting, particularly for small labeled datasets.

### C. LOCAL TRAINING ON LABELED DATA

Once the labeled dataset was preprocessed, each client  $C$  with labeled data initiates local training. This phase aims to learn a strong feature representation and segmentation capability using supervised learning before moving on to semi-supervised training with unlabeled data.

The training begins with feature extraction using a DenseNet-121 encoder, a deep convolutional neural network known for its ability to preserve spatial information through densely connected layers. The extracted feature maps are subsequently passed to a segmentation network, which predicts the target regions within the medical images.

To optimize segmentation, the model employs a composite

loss function, combining cross-entropy loss and dice loss. The cross-entropy loss ensures that the network efficiently learns class probabilities, while dice loss addresses class imbalance, a common issue in medical image segmentation. This combination helps refine boundary delineation and improves overall segmentation.

Mathematically, the cross entropy loss function is defined as:

$$\mathcal{L}_{CE}^{(i)} = -\frac{1}{m} \sum_{j=1}^m \sum_{c=1}^T y_{j,c} \log(f_{\theta}^{(i)}(x_j)_c) \quad (1)$$

Where,  $T$  is number of classes ( $T = 8$ ), and  $x_j$  is  $j$ -th input sample.  $y_{j,c}$  is the one-hot encoded ground truth label for sample  $j$  in class  $c$ .  $f_{\theta}^{(i)}(x_j)_c$  is the predicted probability for class  $c$  in the Equation (1).

The overall loss function, incorporating both cross-entropy and dice loss, is expressed as:

$$\mathcal{L} = \lambda_1 \mathcal{L}_{CE} + \lambda_2 \mathcal{L}_{Dice} \quad (2)$$

Where  $\mathcal{L}_{CE}$  is the cross-entropy loss,  $\mathcal{L}_{Dice}$  is the Dice loss, and  $\lambda_1, \lambda_2$  are weighting parameters controlling the contribution of each loss term. The Dice loss is particularly beneficial for medical image segmentation, as it enhances sensitivity to small structures and improves the model's robustness against class imbalance.

Training occurs locally on each client for multiple epochs using their respective labeled data subsets. This ensures that each client learns distinct yet relevant representations based on their specific data distributions. The resulting model weights are stored as:

$$\text{client}_i.\text{pth}, i \in \{1, 2, 3, 4\} \quad (3)$$

These weights store important patterns for cancer classification while keeping patient data private. Since the data stays on each client's device, this federated learning method ensures privacy and security.

After completing the local training phase, each client shares its trained model weights with the central server for aggregation. This step ensures that knowledge from distributed data sources is effectively combined while maintaining data privacy. The aggregation follows the Federated Averaging (FedAvg) algorithm, defined as:

$$\theta_{global}^L = \sum_{i=1}^N \frac{m_i}{M} \theta_i \quad (4)$$

Where,  $\theta_i$  represents the model parameters of the  $i$ -th client,  $m_i$  is the dataset size of the client, and  $M$  is the total data size across all clients as shown in the algorithm 1. This aggregated model  $M^L$  is then distributed to all clients for further training on unlabeled data.

The resulting aggregated model, denoted as  $M^L$ , encapsulates the learned representations from all clients trained on labeled data. This global model is then distributed back

to clients as an initial model for SSL on unlabeled data, facilitating further refinement through knowledge transfer.

#### D. PSEUDO-LABELING FOR UNLABELED DATA

In FL, acquiring fully labeled datasets for all clients is often infeasible due to privacy concerns and annotation costs. To address this, a SSL approach is employed, where pseudo-labeling is used to leverage the information from unlabeled data while maintaining data privacy.

The pre-trained model  $M^L$ , obtained from the initial training of the labeled data, is used to create pseudo-labels for the unlabeled dataset of each client  $C$ . This process enables clients to extract meaningful patterns from their local data without explicit manual annotations, improving the robustness of the federated model.

Initially, features are extracted using the same encoder model deployed for labeled data. The segmentation model uses these features to make predictions. However, instead of directly using all predictions as pseudo-labels, a CBDPL mechanism is applied to minimize noisy labels.

For each unlabeled sample  $u_k \in D_i^U$ , the trained model outputs softmax probabilities  $f_\theta^{(i)}(u_k)_c$  represents the predicted probability for class  $c$ . The pseudo-label  $\hat{y}_k$  is assigned based on the highest probability exceeding a predefined confidence threshold  $\tau$ , ensuring only high-confidence predictions contribute to the training process:

$$\hat{y}_k = \arg \max f_\theta^{(i)}(u_k), \text{ if } \max f_\theta^{(i)}(u_k) > \tau \quad (5)$$

Here,  $\tau$  is set to 0.8 to filter out uncertain predictions, reducing the impact of incorrect pseudo-labels, it follows the algorithm 1

Once pseudo-labels are generated, they are merged with the original labeled dataset to create an enriched training set:

$$D_i^{\text{combined}} = D_i^{(L)} \cup D_i^{(\hat{U})} \quad (6)$$

To ensure balanced learning, the model is trained using a hybrid loss function that integrates both supervised and semi-supervised objectives:

$$\mathcal{L}^{(i)}_{\text{Semi}} = \mathcal{L}^{(i)}_{\text{CE}} + \lambda \cdot \frac{1}{n} \sum_{k=1}^n \hat{y}_k \log(f^{(i)}\theta(u_k)) \quad (7)$$

Here,  $n$  is the number of pseudo-labeled samples,  $\lambda$  is a weighting parameter (default:  $\lambda = 1$ ) that controls the influence of pseudo-labeled data in training.

One of the key challenges in pseudo-labeling is the risk of propagating incorrect labels, especially in complex medical imaging tasks where misclassification could lead to severe consequences. To mitigate this:

**Confidence-Based Filtering:** The threshold  $\tau$  ensures that only reliable pseudo-labels are used.

$$\hat{y}_k = \begin{cases} \arg \max f_\theta^{(i)}(u_k), & \text{if } \max f_\theta^{(i)}(u_k) > \tau \text{ discard,} \\ \text{otherwise} & \end{cases} \quad (8)$$

**Gradual Label Integration:** Instead of using all pseudo-labels at once, incremental updates refine the model progressively. Given a batch of pseudo-labeled samples  $\hat{D}_i$ , at each round  $t$ , only a subset  $S_t \subseteq \hat{D}_i$  is incorporated into training:

$$S_t = u_k \in \hat{D}_i \mid \max f^{(i)}\theta(u_k) > \tau_t \quad (9)$$

where  $\tau_t$  is an adaptive threshold that increases over training rounds to incorporate more pseudo-labels gradually.

**Class Distribution Awareness:** To prevent bias toward dominant categories, class-balanced weighting is applied, where the contribution of each class  $c$  in the pseudo-label set is weighted by its inverse frequency  $w_c$ :

$$w_c = \frac{1}{\sqrt{\text{count}(c) + \epsilon}} \quad (10)$$

$$\mathcal{L}^{(i)}_{\text{Balanced}} = \sum c w_c \cdot \mathcal{L}_{\text{PL},c}^{(i)} \quad (11)$$

Where  $\epsilon$  is a small constant to avoid division by zero.

After training on the combined dataset, each client updates its local model and shares it with the central server. The FedAvg algorithm is then used to aggregate these models into a global model  $M^U$ , which integrates knowledge from labeled and pseudo-labeled data across all participating clients. The updated model  $M^U$  is then redistributed to clients for further rounds of federated training, progressively improving segmentation while maintaining privacy.

By dynamically incorporating high-confidence pseudo-labels, this approach effectively bridges the gap between labeled and unlabeled data, enhancing model generalization without requiring additional manual annotations.

#### E. FEDERATED LEARNING FRAMEWORK

FL allows multiple clients to train a shared model without exchanging raw medical images. Each client trains a model on its local data and only shares model updates with a central server. This helps protect patient privacy while enabling collaborative learning across different locations.

To integrate knowledge from labeled and pseudo-labeled datasets, the central server aggregates the two models,  $M^L$  from labeled data training and  $M^U$  from pseudo-labeled-based training using the FedAvg algorithm:

$$\theta_{\text{final}} = \text{FedAvg}(\theta_{M^L}, \theta_{M^U}) \quad (12)$$

where  $\theta_{M^L}$  and  $\theta_{M^U}$  represent the learned parameters from labeled and pseudo-labeled training, respectively. FedAvg ensures that the final model leverages both supervised learning and SSL, improving its generalization to unseen data, as shown in the algorithm 1.

This aggregation process integrates information from labeled and pseudo-labeled datasets, enhancing the model's generalization and segmentation. By using both types of training data, the final model becomes more reliable and performs well across different client datasets.

Once aggregated, the final model is stored on the central server and made available for download by participating

---

**Algorithm 1** Federated Semi-Supervised Segmentation with CBDPL

**Input:** Pre-trained DenseNet-121, client datasets  $\{D_1, D_2, \dots, D_m\}$ , unlabeled dataset  $D_U$ , rounds  $\mathcal{R}$ , confidence threshold  $\tau$ , device  $d$ , save path  $\mathcal{P}$ .

**Output:** Optimized global model  $\theta^*$ , pseudo-labeled dataset  $D_U^{\sim}$ , saved weights.

```
1: Initialize Model:
2:  $f_{\theta} \leftarrow$  DenseNet-121 (pretrained)
3: Adjust output layer:  $Linear(1024, 2)$ 
4: Define loss:  $\mathcal{L} = CrossEntropyLoss(y, \hat{y})$ 
5: Set optimizer:  $\mathcal{O} = Adam(f_{\theta}.parameters(), lr = 0.001)$ 
6: Federated Training:
7: for  $r = 1$  to  $\mathcal{R}$  do
8:    $client\_models \leftarrow \emptyset$ 
9:   for each client  $C_i \in \{D_1, \dots, D_m\}$  do
10:     $f_{\theta_i} \leftarrow f_{\theta}$ 
11:     $f_{\theta_i} \leftarrow ClientUpdate(f_{\theta_i}, D_i)$ 
12:    Append  $f_{\theta_i}$  to  $client\_models$ 
13:   end for
14:   Aggregate models:  $\theta \leftarrow FedAvg(\{f_{\theta_1}, \dots, f_{\theta_m}\})$ 
15:   Update global model:  $f_{\theta} \leftarrow \theta$ 
16: end for
17: Pseudo-Label Generation:
18:  $D_U^{\sim} \leftarrow \emptyset$ , set  $f_{\theta}$  to eval mode.
19: for each batch  $x_i \in D_U$  do
20:    $y_i = f_{\theta}(x_i)$ 
21:    $c_i = \max(Softmax(y_i))$ 
22:   if  $c_i \geq \tau$  then
23:     Assign pseudo-label:  $\hat{y}_i = \arg \max(y_i)$ 
24:     Add  $(x_i, \hat{y}_i)$  to  $D_U^{\sim}$ 
25:   end if
26: end for
27: Save Model and Predictions:
28: Visualize predictions for  $D_U^{\sim}$ 
29: Save model:  $torch.save(f_{\theta}.state\_dict(), \mathcal{P})$ 
30: Output:  $\theta^*$ ,  $D_U^{\sim}$ , saved model.
```

---

clients. This enables each client to benefit from the collective learning process while preserving data privacy. Clients can further fine-tune the global model on their local datasets if necessary, ensuring adaptability to institution-specific variations. The FL framework thus facilitates collaborative model training while maintaining data security and confidentiality.

### 1) Model Upload to S3

To facilitate the federated learning workflow, trained models incorporating labeled and pseudo-labeled datasets are securely uploaded to a centralized AWS S3 bucket using the *upload\_model.py* script. This process ensures efficient model storage and centralized accessibility while upholding data privacy, as no raw data is exchanged between clients.

The upload procedure consists of saving the locally trained model weights after each training round and transferring

them to the designated S3 bucket. This enables seamless integration with the server-side aggregation process, allowing the central server to retrieve and combine model updates efficiently. By leveraging AWS S3, the framework ensures scalable, secure, and reliable model storage, supporting continuous FL cycles.

### 2) Global Model Distribution:

After computing the global model parameters  $\theta_{global}$ , the server distributes the updated model to all participating clients. This step ensures that each client benefits from the collective knowledge gathered during the FL process, allowing them to continue refining their models with the newly aggregated parameters. By incorporating updates from multiple clients, the global model becomes more robust and adaptable to diverse data distributions.

Upon receiving  $\theta_{global}$ , each client resumes local training using its labeled and pseudo-labeled data. This enables further fine-tuning based on local dataset characteristics while leveraging the improvements from the global update. The iterative process comprising local training, model uploads, federated averaging, and global model distribution continues for multiple rounds, progressively enhancing the model's performance. Over successive iterations, this approach leads to a well-generalized model capable of accurate segmentation across various client datasets, improving both convergence and overall predictive accuracy.

## IV. EVALUATION

FL enables collaborative training of deep learning models across distributed medical datasets while preserving data privacy. This evaluation section assesses the effectiveness of our FL framework in segmenting different tumors from medical images.

### A. EXPERIMENT SETUP

The experimental setup for this study is designed to simulate an FL environment on a single machine while ensuring the model's scalability and efficiency. The system is configured on an HP Windows laptop, utilizing VS Code as the primary Integrated Development Environment (IDE) for developing and testing the model. This setup provides a development platform that is familiar and flexible, enabling rapid iteration and debugging during the experimentation phase.

The model training and data storage are handled through AWS services, with a specific focus on Amazon S3 buckets. These cloud services ensure secure, scalable storage of medical imaging datasets, enabling efficient data access and management across federated clients. By using Amazon S3, the project avoids the risk of storing sensitive medical data on local machines, which is essential for preserving privacy in medical image analysis.

The FL process involves simulating multiple clients on the local machine. These simulated clients represent different data nodes in a distributed environment, each responsible for training the model on a subset of the dataset. In a typical FL

setup, these clients could be separate devices (e.g., hospitals or clinics), but for this experiment, they are virtualized on the local laptop for ease of setup and control.

The local machine acts as the central server, which combines the model updates from the clients and coordinates the FL rounds. This server aggregates the model updates after each round of local training and distributes the updated model to the clients for the next round. The process ensures that the clients learn collaboratively without sharing their raw data, thus preserving data privacy.

The training pipeline is implemented in Python, utilizing popular libraries like PyTorch for model development and NumPy for numerical operations. PyTorch's flexibility and scalability make it an ideal choice for implementing the federated learning algorithm. MATLAB is employed for visualization and result interpretation, providing powerful tools for graphing performance metrics, such as accuracy, precision, recall, and Dice coefficient. MATLAB's robust plotting capabilities help to visualize model performance over time and interpret complex results.

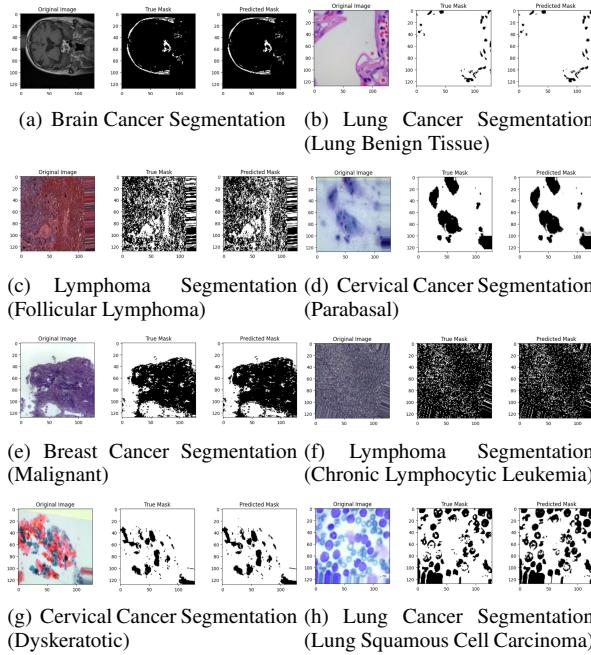


FIGURE 3. Visualization of segmentation results for cancer images.

## B. EVALUATION RESULTS

The dataset comprises labeled and unlabeled cancer CT images. The labeled dataset was used for supervised learning, while pseudo-labeling techniques were applied to the unlabeled dataset for semi-supervised training. Data splitting followed a ratio of 70:30 for training and testing across four federated clients.

Figure 3 presents a visualization of segmentation results for various types of cancer, including brain, cervical, lymphoma, breast, and kidney cancer. Each row corresponds to a different cancer type, showing the original medical

image, the ground truth segmentation mask, and the predicted segmentation mask generated by the model. The results highlight the model's ability to segment tumor regions, with varying levels of accuracy across different cancer types. Some segmentation demonstrates close alignment with ground truth masks. These visual comparisons provide qualitative insights into the effectiveness of the proposed federated learning framework in medical image segmentation.

## C. RESULTS ON LABELED DATA

In the first phase, each client trained a model using labeled data to learn to segment tumors accurately. The key part of this training was DenseNet-121, which acted as the feature extractor. DenseNet-121 is a special neural network that connects every layer to all previous layers. This helps the model reuse important features, improve learning, and capture fine details in medical images.

Each client processed labeled medical images through DenseNet-121, which extracted important features and passed them to the segmentation network. The segmentation model then predicted tumor regions. To improve accuracy, we used two loss functions: cross-entropy loss to classify different regions correctly and dice loss to improve segmentation, especially for small tumor areas.

After training, the model losses for each client are shown in the table 3.

Client ID	Loss	Accuracy	Dice Coefficient
Client 1	0.1306	0.89	0.875
Client 2	0.0371	0.96	0.945
Client 3	0.0312	0.97	0.955
Client 4	0.0353	0.96	0.950

TABLE 2. Different performance metrics after training labeled dataset

These results show that each client successfully learned how to segment tumors. DenseNet-121 helped extract meaningful features to make accurate predictions. Once local training was complete, the models from all clients were combined using FedAvg to create a single global model.

This global model  $M^L$  was then shared with all clients, providing a strong starting point for the next phase, training on unlabeled data.

## D. RESULTS ON SEMI-SUPERVISED LEARNING WITH UNLABELED DATA

SSL was applied to the unlabeled data to improve model performance. CBDPL was applied as the key technique. First, the model was trained using only labeled data, then this trained model was used to predict labels for the unlabeled images. If the model was highly confident in a prediction, it was treated as a pseudo-label and was added to the training data. This allowed the model to learn from more examples without requiring manual labeling.

To ensure accuracy, a Confidence Threshold was set, which meant only predictions above a certain confidence

level were used as Pseudo-labels. This prevented the model from learning incorrect labels, making training more stable. After adding these pseudo-labeled images to the dataset, the model was retrained to refine its performance.

As seen in Figure 4, after applying pseudo-labeling, the model's average loss was 0.4869, and its average accuracy was 0.875, which is close to what we achieved using the fully labeled data. This proves that SSL helps models learn better, even when labeled data is limited. The final Pseudo-label loss was 0.3048 and the accuracy was 0.927.

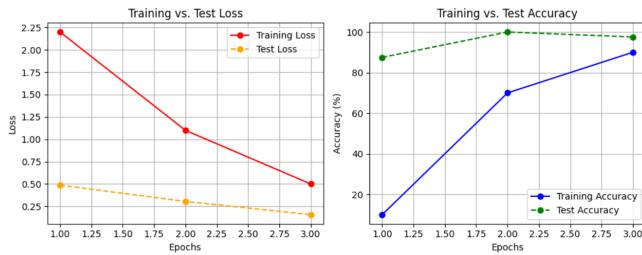


FIGURE 4. The training and test loss and accuracy over time

In Figure 4, the training and test loss over time were tracked. The left graph with the training loss (red solid line) and test loss (orange dashed line) shows the loss reducing steadily, meaning the model was learning well. The right graph shows accuracy trends, where both training accuracy (blue solid line) and test accuracy (green dashed line) improved, confirming that pseudo-labeling helped the model learn effectively.

To better understand how the model generated pseudo-labels, the visualization of the confidence scores in figure 5 can be used. The left plot (a) presents a line graph that represents the confidence scores across different samples. Each point represents a sample, and its confidence score indicates how certain the model is about its pseudo-label. The fluctuation suggests that some samples have highly confident predictions, while others are less certain, likely due to variations in image quality or complexity. The right plot (b) visualizes the overall confidence score distribution using a histogram with a density curve. The histogram represents how frequently different confidence scores appear across the dataset, while the density curve provides a smooth estimate of the distribution. A well-balanced confidence distribution ensures that pseudo-labeling improves model performance without introducing excessive noise.

By analyzing these confidence scores, the optimal threshold for selecting pseudo-labels can be determined, ensuring that only high-certainty predictions contribute to the training. The results confirm that pseudo-labeling significantly enhances model robustness, reducing loss while improving accuracy.

#### E. COMPARISON WITHIN OUR MODEL

The performance of the CBDPL model was evaluated across multiple clients to see how well it worked in a distributed setup. Each client trained the model on its data and then

shared updates to improve the global model. This method keeps data private while allowing all clients to learn together. The evaluation focused on training loss, accuracy, and consistency of learning across different clients.

Figure 6 shows how the training loss (left) and training accuracy (right) change over time for different clients using FL. All clients start with a relatively high loss in training loss, which decreases as training progresses. Most clients have much lower loss by the third epoch, meaning the model has learned well. However, Client 1 has a slightly higher loss than the others, which may be due to differences in its data. In Training accuracy, the accuracy improves steadily for all clients, with most reaching around 98% by the third epoch. Client 1, however, shows a smaller improvement, possibly due to differences in data distribution.

The accuracy and loss of each client after 10 rounds of training were measured, as shown in Table 4

Client ID	Rounds	Accuracy	Loss
Client 1	10	96.93%	0.1372
Client 2	10	97.40%	0.0935
Client 3	10	97.45%	0.0905
Client 4	10	97.08%	0.0974

TABLE 3. Global Model Performance

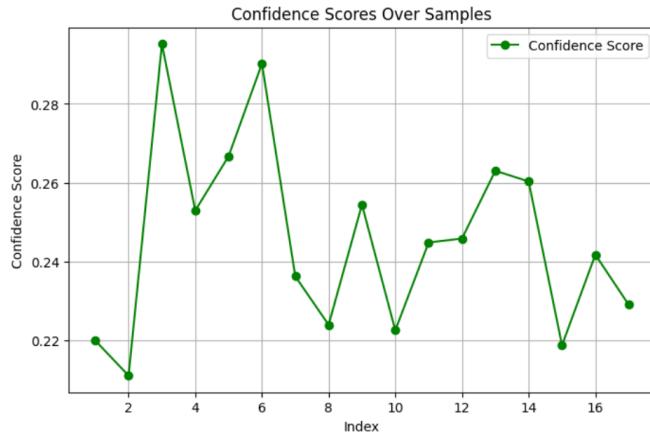
The FedAvg global model achieved a loss of 0.1567 and an accuracy of 0.975.

The model was further evaluated using 5-fold cross-validation. The dataset was divided into five equal subsets (folds), and the model was trained on four folds while the remaining fold was used for testing in each iteration. This process was repeated for all five folds, ensuring that every subset was used for testing once. The average accuracy and loss across all folds were computed to assess model stability. The results are summarized in Table 4. The mean accuracy of 0.9755 and an average loss of 0.1547 indicate that the model achieved stable performance across different training subsets. The low standard deviation further confirms the robustness of the model.

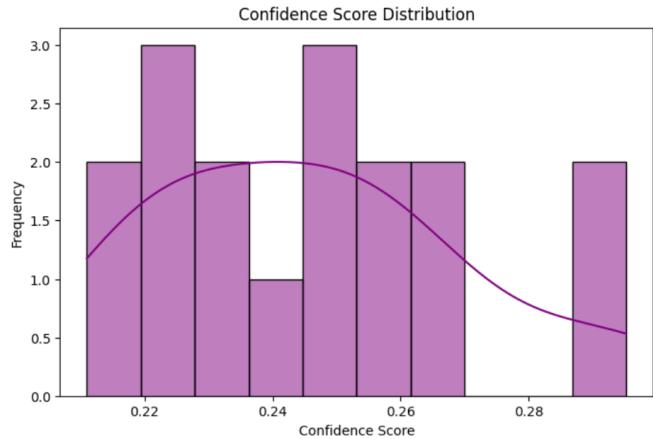
TABLE 4. 5-Fold Cross-Validation Results

Fold	Accuracy	Loss
Fold 1	0.9745	0.1623
Fold 2	0.9762	0.1508
Fold 3	0.9748	0.1542
Fold 4	0.9771	0.1489
Fold 5	0.9750	0.1573
<b>Mean</b>	<b>0.9755</b>	<b>0.1547</b>
<b>Standard Deviation</b>	<b>0.10%</b>	<b>0.0051</b>

These results supports the effectiveness of our proposed CBDPL approach, achieving high accuracy while maintaining privacy by ensuring that data remains distributed across clients.



(a) Confidence Scores Over Samples



(b) Confidence Score Distribution

FIGURE 5. Visualization of Confidence Scores after Pseudo-Labeling

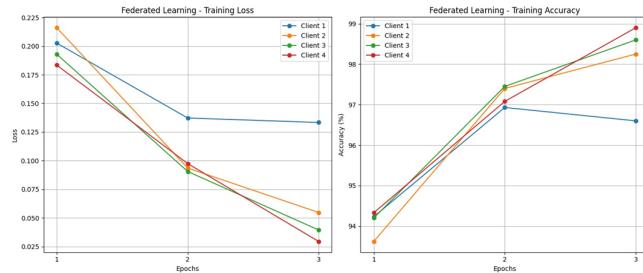


FIGURE 6. Comparison of Training loss and accuracy among the clients

## F. COMPARISON OF RESULTS WITH EXISTING STUDIES

To evaluate the effectiveness of the proposed FL framework, we compared our results with existing studies on medical image segmentation. The comparison considers segmentation accuracy, loss, and learning methodologies such as centralized, federated, and SSL approaches.

As seen in Table 5, our federated SSL approach outperformed existing FL methods in terms of accuracy and loss. The inclusion of confidence-based pseudo-labeling contributed to better generalization, particularly for unlabeled medical images.

To contextualize the performance of our proposed segmentation framework, we compared the FedAvg algorithm, used in our implementation, with other prominent FL strategies. As shown in Table 6, FedAvg achieved a strong global accuracy of 97.55% with a corresponding loss of 0.1567, indicating a positive convergence of the model between the participating clients. Although alternative methods such as FedMA showed slightly improved accuracy (97.45%) and reduced loss, they often require more complex optimization routines or layer-wise model alignment, increasing computational overhead. FedSGD and FedProx, although useful in specific non-iid settings, did not outperform FedAvg in our evaluation. Given its simplicity, communication efficiency, and robust performance, FedAvg was selected as the aggrega-

TABLE 5. Comparison of Results with Existing Studies

Study	Method	Dataset	Acc. (%)
Our Model	CBDPL in FL	Multi Cancer	97.55
Albalawi et al. [1]	FL + Transfer Learning	Brain Tumor	96.45
Shahzad et al. [2]	Federated Learning	Spine Surgery	95.80
Yan et al. [8]	Variation-Aware FL	Multi-Source	96.20
Ullah et al. [9]	Scalable FL	Medical Imaging	97.10
Yue et al. [10]	Specificity-Aware FL	Imbalanced Images	96.75
Lei et al. [11]	FL Domain Adaptation	Alzheimer's	96.40
Yan et al. [12]	Self-Supervised FL	Medical Imaging	96.50
Rajagopal et al. [14]	FL for MRI Detection	Prostate Cancer	96.30
Hossain et al. [18]	Collaborative FL	Lung/Colon Cancer	97.25

tion method for our system. This choice supports both scalability and practical deployment in real-world medical settings where data privacy and resource constraints are critical.

FL Method	Accuracy	Loss
FedAvg (our model)	97.55%	0.1567
FedSGD	96.85%	0.1682
FedProx	97.22%	0.1478
FedMA	97.45%	0.1439
EVIL	96.87%	0.1812

TABLE 6. Global Model Performance Comparison Across Federated Learning Methods

The EVIL method by Yingyu et al. offers a valuable approach to semi-supervised medical image segmentation by incorporating uncertainty through Dempster-Shafer Theory, which effectively handles ambiguous predictions. However, in our FL setting, our CBDPL method shows improved performance by dynamically selecting only high-confidence pseudo-labels. This targeted approach allows for more effi-

cient use of unlabeled data without discarding useful information, while also being computationally lighter and better suited for federated environments where clients train locally on diverse data. These design choices help our method achieve higher accuracy and faster convergence, making it well adapted to the practical challenges of privacy-preserving multi-institution medical image segmentation.

These results show that FL can improve medical image segmentation, especially when combined with SSL methods. The privacy-preserving advantages of FL make it a viable solution for collaborative medical research without compromising performance.

## V. DISCUSSION

The evaluation results from this study underscore the effectiveness of FL in achieving high performance for medical image segmentation tasks. The experimental setup, leveraging AWS for secure data storage and a locally simulated federated environment, enabled a privacy-preserving framework that maintained model performance comparable to centralized approaches.

On the labeled dataset, the model demonstrated impressive segmentation accuracy across all clients. Loss values for clients ranged from 0.0312 to 0.1306, indicating consistent performance despite variations in local data distributions. These results validate the capability of FL to achieve robust outcomes in a decentralized setting, even when individual client datasets are small and heterogeneous. Furthermore, the FedAvg aggregated model achieved a loss of 0.1567 and an accuracy of 97.55%, closely aligning with the centralized model's benchmark performance. This demonstrates that federated learning can deliver high-quality results without requiring centralized access to sensitive data.

The integration of SSL significantly enhanced the system's ability to utilize unlabeled data. By employing confidence-based pseudo-labeling, the model achieved an average accuracy of 87.5% on semi-supervised tasks, with a notable loss reduction to 0.3048 for the pseudo-labeled data. These findings highlight the potential of semi-supervised federated learning to generalize effectively to unseen data, even with a substantial portion (70%) of the dataset being unlabeled. The confidence scores, ranging between 0.2187 and 0.2952, ensured that only high-quality pseudo-labels were included in training, thereby improving overall robustness.

Comparison with the centralized model reinforces the feasibility of FL for medical imaging applications. While the centralized model leveraged aggregated datasets to achieve slightly better performance, the federated approach achieved nearly equivalent results (loss: 0.1567, accuracy: 97.55%) while preserving data privacy. For instance, Client 3 achieved a loss of 0.0905 and an accuracy of 97.45% during federated training, further substantiating the model's ability to handle data decentralization effectively.

These results collectively demonstrate that FL, augmented by semi-supervised techniques, offers a scalable and privacy-preserving solution for medical image segmentation. How-

ever, observed variations in client-specific losses suggest the potential for incorporating client-specific optimizations or personalization strategies to further improve outcomes. Moreover, fine-tuning the pseudo-labeling algorithm could enhance model performance in semi-supervised scenarios.

## VI. CONCLUSION

This study demonstrates the effectiveness of FL for cancer segmentation, achieving high accuracy while preserving patient privacy. The proposed approach, combining FedAvg with SSL and confidence-based pseudo-labeling, enables robust training on partially labeled datasets. Experimental results show that federated models achieve accuracy comparable to centralized models, with an aggregated accuracy of 97.55% and improved generalization to unlabeled data. While the framework successfully balances privacy and performance, future work could refine pseudo-labeling strategies and address client heterogeneity for broader applicability. This research underscores the potential of FL in advancing secure and scalable medical image analysis.

## REFERENCES

- [1] Eid Albalawi, Mahesh T R, Arastu Thakur, V Vinod Kumar, Muskan Gupta, Surbhi Bhatia Khan, and Ahlam Almusharraf. Integrated approach of federated learning with transfer learning for classification and diagnosis of brain tumor. *BMC Med. Imaging*, 24(1):110, May 2024.
- [2] Sunder Ali Khowaja, Kapal Dev, Syed Muhammad Anwar, and Marius George Linguraru. Selffed: Self-supervised federated learning for data heterogeneity and label scarcity in medical images. *Expert Systems with Applications*, 261:125493, February 2025.
- [3] Anshu Ankolekar, Sebastian Boie, Maryam Abdollahyan, Emanuela Gadaleta, Seyed Alireza Hasheminasab, Guang Yang, Charles Beauville, Nikolaos Dikaios, George Anthony Kastis, Michael Bussmann, Sara Khalid, Hagen Kruger, Philippe Lambin, and Giorgos Papanastasiou. Advancing oncology with federated learning: transcending boundaries in breast, lung, and prostate cancer. a systematic review, 2024.
- [4] Siguang Chen, Tong Jin, Yanyan Xia, and Xue Li. Metadata and image features co-aware semi-supervised vertical federated learning with attention mechanism. *IEEE Transactions on Vehicular Technology*, 73(2):2520–2532, 2024.
- [5] Yingyu Chen, Ziyuan Yang, Chenyu Shen, Zhiwen Wang, Zhongzhou Zhang, Yang Qin, Xin Wei, Jingfeng Lu, Yan Liu, and Yi Zhang. Evidence-based uncertainty-aware semi-supervised medical image segmentation. *Computers in Biology and Medicine*, 170:108004, 2024.
- [6] Rémi Gosselin, Loïc Vieu, Faiza Loukil, and Alexandre Benoit. Privacy and security in federated learning: A survey. *Applied Sciences*, 12(19), 2022.
- [7] Qi Guo, Yong Qi, Saiyu Qi, and Di Wu. Dual class-aware contrastive federated semi-supervised learning, 2023.
- [8] Arash Heidari, Danial Javaheri, Shiva Toumaj, Nima Jafari Navimipour, Mahsa Rezaei, and Mehmet Unal. A new lung cancer detection method based on the chest ct images using federated learning and blockchain systems. *Artificial Intelligence in Medicine*, 141:102572, 2023.
- [9] Md. Munawar Hossain, Md. Robiul Islam, Md. Faysal Ahmed, Mominul Ahsan, and Julfikar Haider. A collaborative federated learning framework for lung and colon cancer classifications. *Technologies*, 12(9), 2024.
- [10] Amelia Jiménez-Sánchez, Mickael Tardy, Miguel A. González Ballester, Diana Mateus, and Gemma Piella. Memory-aware curriculum federated learning for breast cancer classification. *Computer Methods and Programs in Biomedicine*, 229:107318, 2023.
- [11] Madhura Joshi, Ankit Pal, and Malaikannan Sankarasubbu. Federated learning for healthcare domain - pipeline, applications and challenges. *ACM Transactions on Computing for Healthcare*, 3(4):1–36, October 2022.

- [12] Fei Kong, Xiyue Wang, Jinxi Xiang, Sen Yang, Xinran Wang, Meng Yue, Jun Zhang, Junhan Zhao, Xiao Han, Yuhan Dong, Biyue Zhu, Fang Wang, and Yueping Liu. Federated attention consistent learning models for prostate cancer diagnosis and gleason grading. *Computational and Structural Biotechnology Journal*, 23:1439–1449, 2024.
- [13] Baiying Lei, Yun Zhu, Enmin Liang, Peng Yang, Shaobin Chen, Huoyou Hu, Haoran Xie, Ziyi Wei, Fei Hao, Xuegang Song, Tianfu Wang, Xiaohua Xiao, Shuqiang Wang, and Hongbin Han. Federated domain adaptation via transformer for multi-site alzheimer's disease diagnosis. *IEEE Transactions on Medical Imaging*, 42(12):3651–3664, 2023.
- [14] Lingxiao Li, Niantao Xie, and Sha Yuan. A federated learning framework for breast cancer histopathological image classification. *Electronics*, 11(22), 2022.
- [15] Jingwei Liu, Jin Zhang, Mian Ahmad Jan, Rong Sun, Lei Liu, Sahil Verma, and Pushpita Chatterjee. A comprehensive privacy-preserving federated learning scheme with secure authentication and aggregation for internet of medical things. *IEEE Journal of Biomedical and Health Informatics*, 28(6):3282–3292, 2024.
- [16] Eric Appiah Mantey, Conghua Zhou, Joseph Henry Anajemba, John Kingsley Arthur, Yasir Hamid, Atif Chowhan, and Obinna Ogbonnia Ottu. Federated learning approach for secured medical recommendation in internet of medical things using homomorphic encryption. *IEEE Journal of Biomedical and Health Informatics*, 28(6):3329–3340, 2024.
- [17] Ashkan Moradi, Fadila Zerka, Joeran Sander Bosma, Derya Yakar, Jeroen Geerdink, Henkjan Huisman, Tone Frost Bathen, and Mattijs Elschot. Federated learning for prostate cancer detection in biparametric MRI: optimization of rounds, epochs, and aggregation strategy. In Weijie Chen and Susan M. Astley, editors, *Medical Imaging 2024: Computer-Aided Diagnosis*, volume 12927, page 129271Q. International Society for Optics and Photonics, SPIE, 2024.
- [18] Junsheng Mu, Michel Kadoch, Tongtong Yuan, Wenzhe Lv, Qiang Liu, and Bohan Li. Explainable federated medical image analysis through causal learning and blockchain. *IEEE Journal of Biomedical and Health Informatics*, 28(6):3206–3218, 2024.
- [19] Obuli Sai Naren. Multi cancer dataset, 2024.
- [20] Liang Qiu, Jierong Cheng, Huxin Gao, Wei Xiong, and Hongliang Ren. Federated semi-supervised learning for medical image segmentation via pseudo-label denoising. *IEEE Journal of Biomedical and Health Informatics*, 27(10):4672–4683, 2023.
- [21] Abhejit Rajagopal, Ekaterina Redekop, Anil Kemissetti, Rushikesh Kulkarni, Steven Raman, Karthik Sarma, Kirti Magudia, Corey W. Arnold, and Peder E.Z. Larson. Federated learning with research prototypes: Application to multi-center mri-based detection of prostate cancer with diverse histopathology. *Academic Radiology*, 30(4):644–657, 2023. Special Issue: Adapting after COVID.
- [22] Nicola Rieke, Jonny Hancock, Wenqi Li, Fausto Milletarì, Holger R Roth, Shadi Albarqouni, Spyridon Bakas, Mathieu N Galtier, Bennett A Landman, Klaus Maier-Hein, Sébastien Ourselin, Micah Sheller, Ronald M Summers, Andrew Trask, Daguang Xu, Maximilian Baust, and M Jorge Cardoso. The future of digital health with federated learning. *NPJ Digit. Med.*, 3(1):119, September 2020.
- [23] Hania Shahzad, Cole Veliky, Hai Le, Sheeraz Qureshi, Frank M Phillips, Yashar Javidan, and Safdar N Khan. Preserving privacy in big data research: the role of federated learning in spine surgery. *Eur. Spine J.*, February 2024.
- [24] Md Fahimuzzman Sohan and Anas Basalamah. A systematic review on federated learning in medical image analysis. *IEEE Access*, 11:28628–28644, 2023.
- [25] Umamaheswaran Subashchandrabose, Rajan John, Usha Veerasamy Ambazhagu, Vinoth Kumar Venkatesan, and Mahesh Thyluru Ramakrishna. Ensemble federated learning approach for diagnostics of multi-order lung cancer. *Diagnostics*, 13(19), 2023.
- [26] Guanqun Sun, Han Shu, Feihe Shao, Teeradaj Racharak, Weikun Kong, Yizhi Pan, Jingjing Dong, Shuang Wang, Le-Minh Nguyen, and Junyi Xin. Fkd-med: Privacy-aware, communication-optimized medical image segmentation via federated learning and model lightweighting through knowledge distillation. *IEEE Access*, 12:33687–33704, 2024.
- [27] Nguyen Truong, Kai Sun, Siyao Wang, Florian Guittot, and Yike Guo. Privacy preservation in federated learning: An insightful survey from the gdpr perspective, 2021.
- [28] Farhan Ullah, Gautam Srivastava, Heng Xiao, Shamsher Ullah, Jerry Chun-Wei Lin, and Yue Zhao. A scalable federated learning approach for collaborative smart healthcare systems with intermittent clients using medical imaging. *IEEE Journal of Biomedical and Health Informatics*, 28(6):3293–3304, 2024.
- [29] Praneeth Vepakomma, Otkrist Gupta, Tristan Swedish, and Ramesh Raskar. Split learning for health: Distributed deep learning without sharing raw patient data, 2018.
- [30] Jiacheng Wang, Yueming Jin, Danail Stoyanov, and Liansheng Wang. Feddp: Dual personalization in federated medical image segmentation. *IEEE Transactions on Medical Imaging*, 43(1):297–308, 2024.
- [31] Pochuan Wang, Chen Shen, Masahiro Oda, Chiou-Shann Fuh, Kensaku Mori, Weichung Wang, and Holger R. Roth. Federated learning with partially labeled data: A conditional distillation approach, 2024.
- [32] Zhoushang Wang, Geng Yang, Hua Dai, and Chunming Rong. Privacy-preserving split learning for large-scaled vision pre-training. *IEEE Transactions on Information Forensics and Security*, 18:1539–1553, 2023.
- [33] Jeffry Wicaksana, Zengqiang Yan, Dong Zhang, Xijie Huang, Huimin Wu, Xin Yang, and Kwang-Ting Cheng. Fedmix: Mixed supervised federated learning for medical image segmentation. *IEEE Transactions on Medical Imaging*, 42(7):1955–1968, 2023.
- [34] Huiyi Wu, Baiming Zhang, Cheng Chen, and Jing Qin. Federated semi-supervised medical image segmentation via prototype-based pseudo-labeling and contrastive learning. *IEEE Transactions on Medical Imaging*, 43(2):649–661, 2024.
- [35] Xuanang Xu, Hannah H. Deng, Jamie Gateno, and Pingkun Yan. Federated multi-organ segmentation with inconsistent labels. *IEEE Transactions on Medical Imaging*, 42(10):2948–2960, 2023.
- [36] Rui Yan, Liangqiong Qu, Qingyue Wei, Shih-Cheng Huang, Liyue Shen, Daniel L. Rubin, Lei Xing, and Yuyin Zhou. Label-efficient self-supervised federated learning for tackling data heterogeneity in medical imaging. *IEEE Transactions on Medical Imaging*, 42(7):1932–1943, 2023.
- [37] Zengqiang Yan, Jeffry Wicaksana, Zhiwei Wang, Xin Yang, and Kwang-Ting Cheng. Variation-aware federated learning with multi-source decentralized medical image data. *IEEE Journal of Biomedical and Health Informatics*, 25(7):2615–2628, 2021.
- [38] Ziyuan Yang, Yingyu Chen, Zhiwen Wang, Hongming Shan, Yang Chen, and Yi Zhang. Patient-level anatomy meets scanning-level physics: Personalized federated low-dose ct denoising empowered by large language model, 2025.
- [39] Ziyuan Yang, Wenjun Xia, Zexin Lu, Yingyu Chen, Xiaoxiao Li, and Yi Zhang. Hypernetwork-based physics-driven personalized federated learning for ct imaging. *IEEE Transactions on Neural Networks and Learning Systems*, 36(2):3136–3150, 2025.
- [40] Guanghui Yue, Peishan Wei, Tianwei Zhou, Youyi Song, Cheng Zhao, Tianfu Wang, and Baiying Lei. Specificity-aware federated learning with dynamic feature fusion network for imbalanced medical image classification. *IEEE Journal of Biomedical and Health Informatics*, 28(11):6373–6383, 2024.
- [41] S. Kevin Zhou, Hayit Greenspan, Christos Davatzikos, James S. Duncan, Bram Van Ginneken, Anant Madabhushi, Jerry L. Prince, Daniel Rueckert, and Ronald M. Summers. A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises. *Proceedings of the IEEE*, 109(5):820–838, 2021.



HEMALATHA K received her M.E. degree in Computer Science and Engineering from Anna University, Chennai, in 2019. She has completed her Ph.D. at the College of Engineering, Guindy, Anna University, Chennai, India. She is currently working as an Assistant Professor at Vellore Institute of Technology, Chennai, India. Her research interests include medical image processing and computer vision.



SHRADHANJALI DAS is currently pursuing a Bachelor of Technology (B.Tech) in Computer Science and Engineering with a specialization in Cyber-Physical Systems at VIT Chennai. She is actively exploring various fields within artificial intelligence, machine learning, and deep learning. Her current work includes federated learning, medical image analysis, and privacy-preserving AI techniques.

• • •